

**UNITED STATES DISTRICT COURT
DISTRICT OF MASSACHUSETTS**

SCANSOFT, INC.,

Plaintiff,

V.

VOICE SIGNAL TECHNOLOGIES, INC.,
LAURENCE S. GILICK, ROBERT S. ROTH,
JONATHAN P. YAMRON and MANFRED G.
GRABHERR,

Defendants.

C. A. No. 04-10353-PBS

**VOICE SIGNAL TECHNOLOGIES'
CLAIM CONSTRUCTION MEMORANDUM
U. S. PATENT 6,501,966**

This memorandum will identify the claim construction issues presented by U.S. Patent No. 6,501,966 (the “966 patent”) and will provide defendant Voice Signal Technologies’ proposed claim constructions. Voice Signal will then demonstrate that these constructions are (1) required by the language of the claims, (2) compelled by the written description of the '966 patent, (3) consistent with the prosecution history of the '966 patent, and (4), in an unusual turn of events, supported by the testimony of the inventors and ScanSoft’s Chief Technology Officer.

PRELIMINARY STATEMENT

This memorandum assumes knowledge of terms and concepts that are described in the tutorial prepared by Charles C. Wooters (“Tutorial”), filed herewith. The Court is urged to read the tutorial before reviewing this memorandum.

THE FACTS

Beginning at least as early as 1980, various approaches to the dialing of telephone calls by the use of spoken words were known. All of these systems permitted a user to dial a call by speaking a telephone number, a destination (*e.g., home*) or a person's name. Tutorial, ¶¶ 14,15. Some voice dialing systems were "speaker dependent" (the system was trained by a single user to recognize names and other words spoken by that user), some were "speaker independent" (the system was designed to recognize a limited set of words spoken by any user), and some included both functionalities. *See* Tutorial, ¶¶ 14, 15, 25.

It was also known that the task of speech recognition was simplified if the user were required to speak words in a particular order -- to use a particular "grammar" or "syntax." The user would first speak a particular word or phrase that, if recognized, would allow the speech recognition system to know what class of words -- what subset of its recognition vocabulary -- would be spoken next. This allowed the system to compare the next spoken word or phrase against a relatively small subset of its total vocabulary. It speeded recognition and improved accuracy. Tutorial, ¶¶ 25(c). In the context of voice dialing, the user might, for example, speak the word *dial* if the user intended thereafter to say a telephone number. The system would then "expect" the user to speak a series of numbers and would compare incoming sounds only against its templates for numbers. Alternatively, the user might say *call*, if a destination to be called (*e.g., home* or *office*) was to be spoken next. The system would then compare the sounds next spoken by the user against its templates for destinations.

These early voice dialing systems further improved performance by providing spoken prompts to the user. Prompts were designed to tell the user what type of word(s) the user should speak next so that the user would (or was more likely to) comply with that grammar used by the system. Voice dialers also allowed the user to verify that the user's spoken words had been

recognized accurately by the device. Thus, for example, a speech recognition system might repeat back to the user the telephone number that the user had spoken. Tutorial, ¶¶ 25(e).

Voice dialing systems were developed for wireline (*i.e.*, conventional) telephones and wireless telephones (car phones, cell phones and other mobile phones). Tutorial, ¶ 4. When a voice dialing capability was incorporated in a mobile phone, words were spoken into the handset's mouthpiece. The sounds that made up these words were converted into electrical signals. The signals were passed directly to, and interpreted by, a speech recognition system "embedded" in the phone. The phone's speech recognition system interpreted the sounds spoken by the user and then instructed the phone's dialing mechanism to dial a telephone number. The output from the phone was the same as if the user had "dialed" a telephone number by pushing keys on the phone's keypad. Voice dialing activity was complete before a cellular network's services were used. Tutorial, ¶ 22. Voice Signal's product -- called V-Suite -- is software that imparts a voice dialing capability of this type to a mobile handset.

As early as 1986, the inventors of the '966 patent developed a voice dialing product for use in a mobile phone. The product allowed users to "dial" calls by speaking a telephone number or by speaking a destination (*e.g.*, *home* or *office*). The product, called VoiceDial, was later sold commercially by Uniden. This product employed a "syntax" of the type described above. A similar product was developed and promoted by Texas Instruments.

The Uniden product required large amounts of (expensive) processing power and computer memory. As a result, the cost of the product -- hundreds of dollars -- exceeded its practical utility. *See* Transcript of the Deposition of Peter Foster, 120.¹ In the early 1990's, the inventors of the '966 patent had the insight that enormous processing power and memory

¹ Excerpts from deposition testimony that are cited in this brief are attached as exhibits to the Declaration of Wendy Plotkin filed herewith. Excerpts from the deposition of Peter Foster can be found at Exhibit 1, excerpts from the deposition of Thomas Schalk at Exhibit 2, and excerpts from the deposition of Michael Philips at Exhibit 3.

capacity could be brought to bear if voice dialing were to be a service provided by a telephone network using a centrally-located voice dialing facility. A voice dialing capability of this type would have a high total cost but, because it could be accessed by thousands of network subscribers, a low per-user cost. Foster Dep., 118-120, 123. The inventors believed that a network-based (as distinguished from cell-phone based) voice dialing system would have substantial commercial value, and they decided to patent it. Foster Dep., 110.

In a network-based system of the type described in the '966 patent, the user's spoken dialing instructions are transmitted over a cellular or other non-wireline network to a central location. They are interpreted, and result in dialing action, by the network at that central location. Network-based voice dialing services confronted at least two challenges that were not issues when the voice recognizer was part of a single mobile phone. The network speech recognition system had to be required to issue prompts to, and interpret words spoken by, many different users at the same time. And, the system had to be able to recognize accurately words -- dialing instructions -- spoken by a user that had degraded in the course of transmission "over the air" to the speech recognizer. Tutorial, ¶ 23.

This background, elaborated in the Tutorial and the inventor depositions described below, provides the context for the '966 patent.

THE '966 PATENT

The '966 patent is entitled *Speech Recognition System For Electronic Switches In A Non-Wireline Communications Network*. It describes a voice dialing system for use at the mobile telephone exchange (MTX) of a cellular or other non-wireline network. Two embodiments are disclosed. In the first, a speech recognition system is physically attached to a cellular switch as an external peripheral device. In the second, the voice dialing mechanism is internal to the

MTX. The patent states that the speech recognition system may be integrated with more than one cellular (or conventional) switch by the use of connecting trunk lines.

Users subscribe to the network speech recognition service described in the '966 patent. They access the service by dialing the telephone number of the network's speech recognition function in the conventional way. They push buttons on a keypad. The network switch then establishes a connection to the network's speech recognition system. Thereafter, the user transmits voice dialing instructions over the network to the speech recognition function. The user's mobile phone merely broadcasts the user's dialing instructions to the central dialing facility. The central facility interprets the user's words and gives dialing instructions directly to the network switch. All users of the network's voice dialing service access the same voice dialing function.

Like prior art systems, the '966 patent describes a voice dialing function that has speaker-independent and speaker-dependent capabilities. Tutorial, ¶ 29. Like prior art systems, it describes a system in which the user was required to speak an ordered series of words -- to use a particular syntax -- so that the task of speech recognition would be simplified. And, like prior art systems, the '966 patent describes a variety of prompts directed to users, as well as techniques designed to confirm that the system had recognized accurately the words spoken by the user.

ScanSoft asserts, and by Court Order dated April 14, 2005, is confined to asserting, claims 1-6 of the '966 patent. Claim 1 is as follows:

A speech recognition method for a mobile telecommunications system which includes a voice recognizer capable of recognizing commands and characters received from a mobile telecommunications user, the method comprising the steps of:

receiving a command from the mobile telecommunications user;

determining whether the command is a first or second command type;

if the command is the first command type, collecting digits representing a telephone number to be dialed received from the mobile telecommunications user; and

if the command is the second command type, determining whether a previously stored telephone number is associated with a keyword received from the mobile telecommunications user.

Claims 2-6 depend from Claim 1. They are not infringed if Claim 1 is not infringed.

The following claim construction issues are presented:

- A. What is a speech recognition method *for a mobile telecommunications system . . .*?
- B. What is a *command*?
- C. What is a *first command type*?
- D. What is a *second command type*?
- E. What is the meaning of the phrase *collecting digits representing a telephone number*?

ARGUMENT

I. Relevant Claim Construction Principles.

The Federal Circuit recently set forth the claim construction principles applicable to any case.

In construing claims, the analytical focus must begin and remain centered on the language of the claims themselves, for it is that language that the patentee chose to use to particularly point out and distinctly claim the subject matter which the patentee regards as his invention. . . .

Claim construction requires this court to place the claim language in its proper technological and temporal context. The best tools for this enterprise are the various forms of intrinsic evidence and, when appropriate, extrinsic evidence. . . . The intrinsic evidence, “*i.e.*, the patent itself, including the claims, the specification and, if in evidence, the prosecution history . . . is the most significant

source of the legally operative meaning of disputed claim language.” *Id.* (internal citation omitted); *see also United States v. Adams*, 383 U.S. 39, 49 (1966) (“[I]t is fundamental that claims are to be construed in the light of the specifications and both are to be read with a view to ascertaining the invention.”); *Astrazeneca v. Mutual Pharm. C., Inc.*, 384 F.3d 1333, 1336-37 (Fed. Cir. 2004) (“[E]vidence intrinsic to the patent — particularly the patent’s specification, including the inventors’ statutorily-required written description of the invention — is the primary source for determining claim meaning”).

Gillette Company v. Energizer Holdings, Inc., 405 F.3d 1367 (Fed. Cir. 2005) (citations omitted).

ScanSoft argues at length that a claim term may not be restricted to a particular embodiment in the specification. ScanSoft Mem., pp. 8-9. As a generalization, that is correct. However, “claims are construed in accordance with the rest of the specification of which they are a part, and not contrary to it.” *C. R. Bard v. M3 Systems, Inc.*, 157 F.3d 1340, 1360 (Fed. Cir. 1998). And, “the specification may assist in resolving ambiguity where the ordinary and accustomed meaning of the words used in the claims lack sufficient clarity to permit the scope of the claim to be ascertained from the words alone. *Teleflex Inc. v. Ficosa North America Corp.*, 299 F.3d 1313, 1325 (Fed. Cir. 2002).

II. “A Speech Recognition Method For A Mobile Telecommunications System . . .”

| Claim Limitation | Voice Signal’s Proposed Construction | ScanSoft’s Proposed Construction |
|---|--|---|
| <i>A speech recognition method for a mobile telecommunications system</i> which includes a voice recognizer capable of recognizing commands and characters received from a mobile telecommunications user . . . | A speech recognition method that is used by a mobile telecommunications system, as distinguished from a method used in a mobile (<i>e.g.</i> cellular) telephone. A mobile telecommunications system is a network that connects mobile telecommunications customers each having a mobile telephone to other | A process for recognizing commands and characters spoken by a user. “Mobile telecommunications system” refers to “cellular, satellite (<i>sic</i>) and personal communications network environments.” Col 3, ll. 46-48. A network environment includes all components of a |

| | | |
|--|--|---|
| | telephone customers. A centrally located voice recognizer recognizes words received over a cellular or other non-wireline network. | network, from the central office switching equipment, if any, to the mobile units (e.g. telephones, pdas, and other communications devices) used therein. |
|--|--|---|

This claim limitation requires “a speech recognition method *for* a mobile telecommunications *system*”². A mobile telecommunications system is the hardware, software and business infrastructure that connects a user of a wireless telephone to other telephone users. Cellular network service providers operated mobile telecommunications systems. The word *for* indicates that the claimed method is maintained, or used, by such a wireless telecommunications system.

ScanSoft argues that a cellular handset is part of a cellular network, and that, accordingly, the phrase “A speech recognition method *for* a mobile telecommunications *system* . . .” encompasses a speech recognition functionality that is not maintained or operated by a network, that is entirely internal to a single handset and that merely causes the handset to transmit to the network the same dialing codes as would be transmitted if the handset’s keypad had been used.

This issue is potentially outcome determinative. It is undisputed that the accused product -- Voice Signal’s V-Suite software -- provides a voice dialing function that is entirely internal to

² It is undisputed that this portion of Claim 1 limits the scope of the claim. The steps of the claimed method which follow this preamble are performed by the mobile telecommunications system required by this language. The claimed steps have no meaning if they are read without the context of the preamble. In addition, two of the four claimed steps, and several claims that depend from Claim 1, reference “the mobile telecommunications user.” The only antecedent for these references is the “voice recognizer” in the preamble. Moreover, several dependent claims reference an “apparatus,” the only antecedent basis for which is in the preamble, and Claim 6 references “the mobile telecommunications system,” the claim term that is at issue. *See also* Claim 13. In addition, as will be demonstrated below, the requirement that the speech recognition method is “for a mobile telecommunications network” is what distinguishes Claim 1 (and other claims) from prior art that was before the Patent Office. *NTP, Inc. v. Research In Motion, Ltd.*, 392 F.3d 1336, 1359 (Fed. Cir. 2004) (“Because the limitations of claim 1 . . . derive their antecedent basis from the claim 1 preamble and are necessary to provide context for the claim limitations, the use of these limitations in the preamble limits the claim”).

an individual handset.³ Hearing Transcript, March 16, 2005, p. 8. If Voice Signal's claim construction is adopted, ScanSoft's patent infringement claims will be dismissed.

A. The Written Description.

Claim 1 requires a speech recognition method *for a mobile telecommunications system*. The written description defines the term *mobile telecommunications system*. It states that "a cellular mobile communication system connects mobile telecommunications customers, each having a mobile unit, to land-based customers serviced by a telephone network." 3:35-38. It thus distinguishes between that which connects customers on the one hand, and the customers and their individual "mobile units," on the other. Cellular network service providers connect customers; an individual cell phone does not.

The patent then makes clear that the term *mobile telecommunications system* includes any mobile telecommunications network, not just a cellular network: "As used herein, mobile telecommunications system refers to cellular, satellite and personal communications network environments." 3:46-48⁴. A "personal communications network" is "any non-wireline network." 2:2. This broadening language does not suggest that a single cell phone is, or is part of, a mobile telecommunications system. And, even if ScanSoft's definition of a mobile telecommunications system were accepted, it is impossible to conclude that a voice dialing capability in a single handset is a speech recognition method *for* a mobile telecommunications

³ As a result, the precise construction of this claim limitation is not critical. The ultimate question is whether the claim covers a voice dialing capability that is entirely internal to a handset and completes its activity before the handset sends conventional dialing signals to a network.

⁴ ScanSoft quotes this sentence without placing it in the context of the definition that it modifies. ScanSoft Mem. at 11. In context, a mobile telecommunications system is any cellular, satellite or personal communication (*i.e.*, any non-wireline) network environment that connects mobile telecommunications customers, each having a mobile unit, to land-based customers. The connection function, contained in the definition set forth in the specification, is central to the meaning of the claim term a "mobile telecommunications system."

system. A handset-based voice dialer is internal to one cell phone. It causes that phone to transmit conventional dialing codes to the network.

Voice Signal's claim interpretation is supported by virtually the entire written description of the '966 patent. From beginning to end, the patent teaches the use of a network-based voice dialing capability in which spoken dialing instructions are carried "over" the network and received "at" a mobile telephone exchange. The speech recognition system is described in the specification as a "shared" resource closely associated with a network switch.

The Background Of The Invention describes the problem that the investors faced:

While telephone-based speech recognition systems are known, cellular voice dialing *over* a mobile telephone exchange ("MTX") presents significant challenges for two basic reasons. First, the recognition technology must accommodate a tremendous range of both remotely-mounted and hand-held microphone types. Second, the signal may be band-limited and degraded in transmission *to* the MTX *where* the recognition system will be located.

1:26-33 (emphasis supplied).

The Brief Summary Of The Invention identifies the objects of the invention and then explains how those objects are achieved. First, the objects of the invention:

It is therefore [*i.e.*, in view of the problems described above] an object of the present invention to describe an implementation of a speech recognition system in a cellular or personal communications network environment.

It is a further object of the invention to describe a speech recognition system for use *at* a mobile telephone exchange (MTX) of a cellular or personal communications network. The placement of the speech recognition system *at* the MTX *significantly reduces cost and increases reliability by enabling the switch vendor to install and maintain the system in conjunction with the cellular switch*.

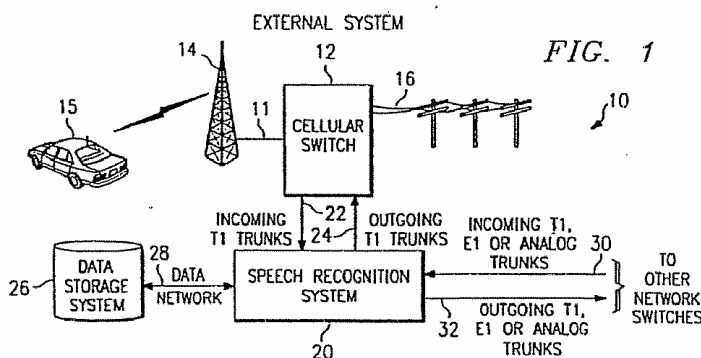
It is another object of the invention to describe a cellular voice dialing system for use *in or in conjunction with* an MTX of a cellular network.

1:44-58 (emphasis supplied). These objects are achieved by a centrally-located, network-based voice dialing capability that is a shared resource:

These and other objects of the invention are provided in an advanced system for the recognizing of spoken commands *over* the cellular telephone or any personal communications (*i.e.*, any non-wireline) network. In the cellular application, for example, a Speech Recognition System interconnects either *internally with or as an external peripheral to a cellular telecommunications MTX switch*. . . . The Speech Recognition System also allows for increased efficiency in the cellular telephone network by integrating with the switch or switches as a shared resource.

1:66-2:12 (emphasis supplied). There is no suggestion whatever that a voice recognition system embedded in a cell phone is located “at” a mobile telecommunications exchange, receives commands “over” a non-wireline network, interconnects internally with, or as an external peripheral to, an MTX switch, or is a “shared resource.” There is, similarly, no suggestion that a handset-based dialing mechanism is consistent with the stated objective of “enabling the switch vendor to install and maintain the voice dialing system.”

The specification proceeds to the Detailed Description of the “invention.” The invention is said to be depicted in Figures 1 and 2, respectively. 3:12, 4:16. Figure 1, reproduced below, “is a block diagram of a cellular telephone network incorporating an external switch-based speech recognition system according to the present invention.” 2:40-42; 3:10-12.

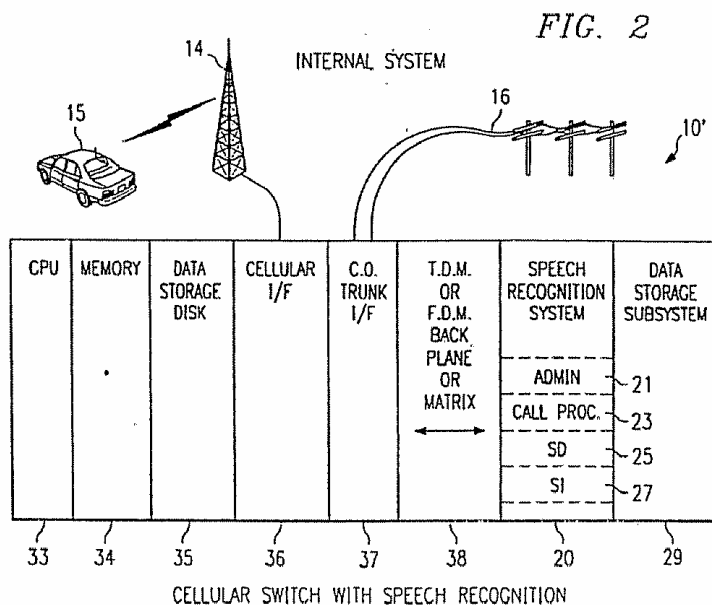


The patent describes the embodiment of Figure 1 as an external peripheral that is physically connected to one or more network switches:

According to one embodiment of the invention as shown in FIG. 1, a Speech Recognition System 20 is connected as an external peripheral to the MTX through a set of preferably digital trunk lines. Set 22 is used for incoming signals and set 24 is used for outgoing signals. . . . The Speech Recognition System 20 may be integrated with one or more switches (whether or not cellular) for use as a shared resource via incoming and outgoing trunk sets 30 and 32.

3:66-4:12. Thus the speech recognition system of the patent can be a “shared resource” connected to any number of wireline or non-wireline networks. 3:20-24, 4:9-12. There is no suggestion that a voice dialer that is internal to a single cell phone -- and that performs each of the steps of the claimed method (if it performs them at all) before transmitting conventional, electronic dialing codes to the network -- is part of the invention.

Figure 2 (reproduced below) depicts a second, and the preferred, embodiment of the “invention.” In this embodiment, the speech recognition system is internal to the MTX.



The patent states:

Referring now to FIG. 2, a block diagram shows the cellular telephone network 10 with the Speech Recognition System 20 interconnected internally to the MTX. This is the preferred embodiment of the invention. The hierarchical architecture of the cellular switch includes the central processing unit 33, memory 34, data storage disk 35, cellular interface 36, central office trunk interface 37 and a back-plane or switching matrix 38.

The Speech Recognition System 20 includes a number of functional subsystems

4:13-20. There is no indication that the invention of the patent is embodied in software that is exclusively located in a mobile telephone that is inside the car 15.⁵ No other embodiment is disclosed.

⁵ ScanSoft argues that Fig. 2 is “a preferred embodiment, not the only one.” ScanSoft Mem. at 14. That is correct. Figure 1 is the other disclosed embodiment.

The patent then explains how a user with a cellular telephone accesses the speech recognition system of the invention. The user manually dials the telephone number of the speech recognition system by pushing keys on a cellular telephone. If the correct number is dialed, the central switch connects the user to the network's speech recognition system, *viz*:

Referring now to FIG. 5, a functional flowchart is provided describing the basic control functions of the Speech Recognition System *according to the present invention*. The routine begins at step 102 when the user dials digits from the cellular telephone. At step 104, a test is performed to determine if a Speech Recognition System access code has been dialed. . . . If the result of the test at step 104 is positive, the routine continues at step 108 during which the switch makes an audio path connection between the user and the Speech Recognition System.

5:59-63 (emphasis supplied). By contrast, the user of a voice dialer embedded in a cell phone does not dial a phone number to get access to a voice dialer, and does not go through the central switch to obtain access to the speech recognition function.

Because the "invention" is a system that is accessed by multiple users, and because it must access data specific to individual users, the invention must be equipped to identify each user. The written description explains that each cell phone is uniquely identified by a mobile identification number, or "MIN." (3:49-53; 6:5-7). When the switch connects a user to the network's speech recognition system, it also transmits the user's cell phone's MIN to the speech recognition system (a step which is, of course, unnecessary where the voice dialing function is "embedded" in a single cell phone). 6:3-7. The network's speech recognition functionality uses the MIN to retrieve from its mass storage device (3:47-48)⁶ data that is unique to a "user who subscribes" to the system's voice dialing "service." 6:8-10. The system of the patent then

⁶ Cell phones do not have "mass storage devices."

retrieves user-specific information, including the language spoken by the user, and the particular telephone numbers that the user has associated with speaker-independent keywords like *home* and *office*. 6:10-15. This step is necessary when a speech recognition system is accessed by persons who speak different languages and who associate different phone numbers with generic location descriptors like *home* or *office*.⁷ It is irrelevant to a voice dialer in a handset.

Once the system of the invention has identified the telephone number that the user wishes to call, that number is transferred to the switch for outdialing. The switch then dials that number. 7:28-33; 8:21-25; 8:52-55. By contrast, a cell phone-based speech recognizer transfers dialing information to the phone's own dialing mechanism, which operates in a conventional manner.

Throughout the written description, users are said to be persons who "subscribe" to a speech recognition service provided by a network. *E.g.* 5:44-45 ("When a user subscribes to the service (*e.g.*, with the MTX service provider)"), 5:51; 6:8-10 ("According to the invention, each user who subscribes to the service will have a prerecorded list of destination numbers. . . ."); 9:52 ("If a subscriber desires to erase a given memory location. . . ."). 9:56; 10:6 ("a test is made . . . to determine whether the subscriber wishes to enroll"); 10:12, 26, 57, 63; 11:1, 2, 3. The owner of a handset which contains an embedded voice dialing capability does not "subscribe" to a voice dialing "service."

The written description then identifies the advantages of the invention. One advantage is that the speech recognition system is able to address multiple users simultaneously: "The invention has multiple advantages over the prior art. . . . Multiple language prompts are spoken from and available simultaneously on multiple ports from a single automated

⁷ A voice dialing system that is internal to a cell phone does not retrieve the language spoken by the user each time a call is made, and does not need to separate the phone number associated by the user with a word like *home* from the phone numbers that other users have associated with the same word.

telecommunications-based system.” 11:20-27. A system that is internal to a single handset could not possibly benefit from such an “advantage.”

A second stated advantage makes it abundantly clear that the alleged invention of the '966 patent is a centrally-located speech recognition function. As noted above, when spoken words are transmitted over the network to a speech recognition facility, they are subject to distortion. Words spoken directly to a voice dialer that is internal to a handset are not. The '966 patent states, in a passage that reveals exactly what it means when it refers to speech recognition “in a non-wireline network,” states:

The invention successfully implements speech recognition in the cellular telephone or personal communications network. Non-wireline networks provide a special challenge to both the recognition algorithm developers as well as the applications developers. The recognition algorithm in conjunction with the system application is insensitive to the radio fading, speech clipping, and speech compression conditions that occur in a non-wireline network. In addition, the recognition algorithm accommodates conditions found in the standard switched network. The invention provides a means of accurately recognizing speech that has limited distortion due to clipping or fading and provides a means of reprompting the user for input when the speech has become too distorted for accurate recognition.⁸

11:31-45. This passage makes clear that the described “invention” is a centrally-located, shared resource that is designed to interpret speech that has been transmitted “over the air” from a remote location.

Finally, it is fair to summarize the '966 patent by quoting the Abstract, thus choosing the inventor’s own words:

An advanced telecommunications system is provided for the recognizing of spoken commands *over* a cellular telephone, satellite telephone, or personal communications network. In the cellular application, for

⁸ The '966 patent does not describe the “recognition algorithm” referenced in the specification and does not otherwise describe the way in which the voice recognizer of the patent interprets the sounds that are transmitted to it. Tutorial, ¶ 40.

example, a Speech Recognition System interconnects either internally with or as an external peripheral to a cellular telecommunications switch. . . . The Speech Recognition System also *allows for increased efficiency in the cellular telephone network by integrating with the switch or switches as a shared resource.*

(emphasis supplied).⁹

As this Court has said, “[t]he Court must construe the claim in the context of the whole patent, with analysis rooted in the words of the specification.” *VLT Corp. v. Unitrade Corp.*, 130 F. Supp. 178, 188 (D. Mass. 201) (Saris, J.). “The specification is always highly relevant to the claim construction analysis. Usually it is dispositive; it is the single best guide to the meaning of a disputed term.” *Vitronics Corp. v. Conceptronic, Inc.*, 90 F.3d 1576, 1582 (Fed. Cir. 1996). The ultimate question is what does the phrase *a speech recognition method for a mobile communications system* mean? The '966 patent defines a “system” as that which connects mobile “customers.” 3:35-36. ScanSoft argues a speech recognition method *for* a mobile telecommunications *system* includes a speech recognition method that is internal to an individual handset.¹⁰ ScanSoft gives the words of Claim 1, viewed in isolation, an unnatural meaning. Read in the context of the written description, ScanSoft’s claim construction is wholly impermissible. It would not merely encompass subject matter that is not described or suggested

⁹ In the absence of language that even hints that mobile unit-based voice dialers are part of its invention, ScanSoft is driven to relying on unspecific, broadening language of the type that appears routinely in almost every patent disclosure. ScanSoft Mem. at 14. Boilerplate statements of this type may not be used to broaden a claim beyond that which is fairly disclosed. *Borg Warner, Inc. v. New Venture Gear, Inc.*, 237 F. Supp. 919, 948, n.23 (N.D. Ill. 2002).

¹⁰ ScanSoft states that Voice Signal’s proposed claim construction requires that the speech recognition system be at a particular location. ScanSoft Mem. at 12-13. It does not. A single speech recognition system may be shared by more than one switch, but the speech recognition system is always maintained at a central location as a shared resource with which users communicate over the network. It is a network-wide resource.

in the '966 patent; it would cover subject matter that is flatly inconsistent with the invention that is described.¹¹

B. The Prosecution History.

In the course of the prosecution of the applications that led to the '966 patent, a variety of prior art handset-based voice dialing systems were cited to the Patent Office. This art is part of the intrinsic claim construction evidence. *Kumar v. Ovonic Battery Co., Inc.*, 351 F.3d 1364, 1368 (Fed. Cir. 2003). Claims should be interpreted “to preserve, rather than defeat, their validity.” *Eastman Kodak Co. v. Goodyear Tire & Rubber Co.*, 114 F.3d 1547, 1556 (Fed. Cir. 1997). *A fortiori*, claims should not be construed to encompass prior art that was before the Patent Office. It is utterly implausible that the inventors intended to obtain, or that the Examiner intended to grant, claims that cover disclosed prior art.

The prior art that was before the Patent Office during the prosecution of the '966 patent included products that utilized, and publications that described, every element of Claim 1 of the patent *if* “a speech recognition method for a mobile telecommunications system” includes a speech recognition method that is entirely internal to a mobile phone.

Claim 1 of the '966 patent requires speech recognition method that includes four steps:

¹¹ The analysis of the '966 patent's written description set forth above demonstrates that ScanSoft's claim differentiation arguments (Mem. at 12-13) are meritless. Claim 8 requires that the speech recognition apparatus “[be] connected to the mobile telecommunications switch as an external peripheral.” This narrows Claim 1, which also encompassed an apparatus that is “interconnected internally to the MTX,” as well as a recognizer that is an external peripheral. 4:15. Claim 11 requires a “speech recognition apparatus [that is] connected to a non-mobile telecommunications switch . . . [which] communicates with the non-mobile telecommunications switch. . . .” This merely singles out a disclosed alternative embodiment that “[t]he Speech Recognition System 20 may be integrated with one or more switches (whether or not cellular) for use as a shared resource via incoming and outgoing trunk sets 30 and 32 ([of Fig. 1].” 4:9-12. It does not suggest that Claim 1 extend to a voice dialer embedded in a handset.

Finally, ScanSoft notes that other patents that issued from the same priority application contain limitations that require a voice recognition system at a non-wireline telecommunications switching office. However, the claims that it cites contain many other limitations that distinguish them from Claim 1 of the '966 patent. They are, for example, product claims while the '966 patent claims a method. Therefore, as a matter of law, the claim differentiation doctrine is inapplicable. *E.g., Macrovision Corp. v. Dwight Cavendish Developments, Ltd.*, 105 F. Supp. 1070, 1074 (N.D. Cal. 2000).

receiving a command from the mobile telecommunications user;

determining whether the command is a first or second command type;

if the command is the first command type, collecting digits representing a telephone number to be dialed received from the mobile telecommunications user; and

if the command is the second command type, determining whether a previously stored telephone number is associated with a keyword received from the mobile telecommunications user.

The specification discloses a particular user interface. The user speaks one of a defined set of commands: *Dial*, *Call* or *Directory*. 6:24-36, Fig. 5. If the command *Dial* is spoken (and recognized by the speech recognition system), the speaker then speaks (and the voice recognizer collects) the digits of a phone number. The number is repeated back to the user and, if the user confirms that his words have been correctly interpreted, the phone number is transferred to the network switch for dialing. See 6:46-7:32, Fig. 5.

If the command *Call* is given (and recognized), the user may then say a *keyword* -- e.g., *home*, *office* or *information*. 7:34-37. The system determines whether a telephone number has been associated with that keyword and repeats it back to the user. Upon confirmation that the number is correct, it is forwarded to the switch for dialing. 7:60-8:25.

Each of these steps was present in mobile phone-based voice dialing systems that were before the Patent Office.

1. Voice Control Systems/Uniden.

In 1986 -- six years before the filing of the application to which the '966 patent claims priority -- Voice Control Systems described in printed publications a speech recognition system for mobile phones. These publications, one written by Thomas B. Schalk, an inventor of the '966 patent (Tutorial, Ex. C), and the other by his colleague Eugene Helms (Tutorial, Ex. D), were

disclosed to the Patent Office during the '966 prosecution. *See* '966 patent, page 2, col. 2, items 8 and 12.

The articles described a system that allowed the user to “dial” a call by voice. The system was incorporated in a mobile phone. They disclosed the same ordered series of instructions by the user, and same actions by the speech recognition system, that are disclosed and claimed in the '966 patent. The user first said one of three words: *dial* (if the digits of a telephone number were to follow), *recall* (if a two-digit number was to be substituted for a full phone number) or *speed dial* (if a “destination description” like *home* or *office* were to follow). If the word *dial* was spoken and recognized, the speech recognition system thereafter collected the digits of the phone number to be dialed. If the words *speed dial* were spoken, the system “listened for” a “destination descriptor” (*e.g.*, *home*, *office*, *information*, etc.) and determined whether a telephone number had previously been associated with that “destination descriptor.” Tutorial, Ex. [Helms] D, p. 130, col. 1; Schalk Dep., pp. 30-45. It then caused that number to be dialed by the phone’s dialing mechanism.

This system became a commercial product, called VoiceDial, that was sold by Uniden. The VoiceDial Operating Guide (Tutorial, Ex. E) was before the Patent Office. '966 patent, p. 2, col. 2, item 13; Foster Dep., p. 54; Schalk Dep., p. 99. VoiceDial was incorporated in a mobile phone. It used the same type of “syntax” that is described in the Schalk and Helms publications. The user first said *dial* or *call*. If the user said (and the product recognized the word) *dial*, the user then spoke the digits of a telephone number, the product collected those digits, repeated the telephone number back to the user, and, upon approval by the user, transferred the phone number to the mobile phone’s dialer. Tutorial, Ex. E, p. 4. If the user said (and the product recognized the word) *call*, the user would then say one of a series of permitted “descriptive words,”

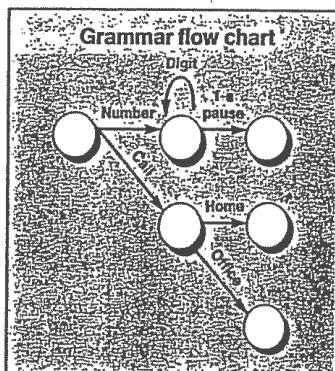
including *home*, *office* and *emergency*. The product would then look up in its memory the phone number that was associated with that word and, upon approval by the user, the relevant phone number would be dialed. Tutorial, Ex. E, p. 10; Foster Dep., pp. 127-128; Schalk Dep., pp. 101-108.

The Uniden VoiceDial performed precisely the same steps, in precisely the same order, as are recited in Claim 1 of the '966 patent. It even used the same command words (*dial* and *call*) and the same keywords as are disclosed in the '966 patent. It was, however, a mobile phone-based, not a network-based, product.

2. Texas Instruments.

An article entitled, "Dialing A Phone By Voice" by Savaraj Pawate and Peter Ehlig appeared in the January 1991 edition of the publication, "Machine Design." Tutorial, Ex. F. This article described a Texas Instruments product, that was "particularly suited for cellular telephones." p. 96, col. 2. Like the Voice Control Systems/Uniden product, a user could "dial" a telephone call by stating a telephone number or a location A (*e.g.*, "home" or "office"). It could also "dial" a call by stating a name ("Harvey"). Tutorial, Ex. F, p. 96, col. 2. The system used an ordered series of instructions and actions, depicted below, that is the same as is disclosed and claimed in the '966 patent. The user either said the word "number" or the word "call." If the user said "number," he was then to speak a telephone number. If the user said "call," he was then to speak the words "home" or "office" or a name (*e.g.*, "John Jones"). p. 97, col. 2.

Other application grammars also are



possible. An application may, for example, require that the speech recognition system recognize names and the word call as in the command call John Jones.

The only distinction between (1) the method claimed in the '966 patent and (2) the methods of the disclosed, prior art Voice Control Systems/Uniden and Texas Instruments products, was that the method claimed in the '966 patent is a “speech recognition method for a mobile telecommunications system” while the prior art disclosed a speech recognition method for a mobile phone. A person of ordinary skill in the art, reading Claim 1 of the '966 patent in the light of the prosecution history of the patent, would necessarily conclude that the phrase “for a mobile telecommunications system” was intended to distinguish the prior art, and, therefore, did not encompass a mobile phone-based speech dialing system.¹²

¹² ScanSoft cites the fact that the PTO Examiner allowed the '966 patent over the prior art Iishi patent, which discloses voice recognition in “wireless telephone apparatus” of some sort. The Examiner concluded that while Iishi described “the number dial as a command type,” it did not disclose “another command type.” ScanSoft argues that the Examiner believed that Iishi was the closest art and states that Iishi describes a “hand set” voice recognizer. ScanSoft Mem. pp. 15-16.

The flaw in this argument, among others, is that nothing in the prosecution history indicates that the Examiner recognized Iishi as handset-based prior art. Indeed, he could not rationally have done so. The prior art Schalk, Helms, Uniden and Texas Instruments references were all before the Examiner. Each was unmistakably hand-set art and each indisputably disclosed exactly what the Examiner said was missing from Iishi.

C. Extrinsic Evidence.

The testimony of the inventors Peter J. Foster and Thomas B. Schalk supports Voice Signal's proposed claim construction. It demonstrates that ScanSoft is now attempting to encompass within the claims of the '966 patent subject matter (a) that the inventors did not intend to claim, and (b) that persons of skill in the art would not believe is claimed in the '966 patent.

Foster and Schalk are paid consultants to ScanSoft. Foster Dep., p. 106; Schalk Dep., p. 12. Their testimony is, of course, extrinsic evidence. Ordinarily, *post hoc* inventor testimony is self-serving and entitled to little weight. However, testimony by an inventor against his interest is "the most persuasive extrinsic evidence." *Evans Medical Ltd. v. American Cyanamid Co.*, 11 F. Supp.2d 338, 350, 351 (S.D.N.Y., 1998) ("the best and most reliable extrinsic evidence"); *aff'd* 1999 WC 594310 (Fed. Cir.) (unpublished); *Perdue Pharma L.P. v. Endo Pharmaceuticals, Inc.*, 2004 WL 26523 *6 (S.D.N.Y. 2004) ("The testimony of the inventors against their own interest is relevant and persuasive to inform the court's claim construction").¹³

Peter Foster was the Chief Executive Officer of Voice Control Systems. Foster Dep., p. 11. He is a named inventor of the '966 patent. Foster testified that Uniden "marketed and sold" the VoiceDial product (Foster Dep. at 54), but that the product was not successful because the hardware needed for VoiceDial made the product too expensive to be attractive to individual consumers. By contrast, when the voice dialing capability became "a piece of the switch or of the network you could spend lots of money on the hardware because it got distributed across tens of thousands of customers." Foster Dep., p. 119. Because a network-based product had real commercial promise, and because, by 1992, the Voice Control Systems/Uniden mobile phone-

¹³ It is notable that ScanSoft repeatedly cites the affidavit of an expert hired for the occasion, but fails to mention the testimony of the inventors themselves.

based product had long since been publicly disclosed (and was thus unpatentable), Foster caused Voice Control Systems to seek patents on the network-based product.¹⁴

Foster explained the practical benefits of the network-based product:

Q. Well, let me ask, were there challenges associated with [the] voice dialing application at the central switch that were not present in the voice dialing application in the handset or in the cellular phone?

A. I'd express it in a different way.

Q. Okay.

A. There was more to do and more you could do and that's the key.

Q. Okay.

A. You didn't see tens of thousands of these NEC phones¹⁵ running around. And they didn't meet the fundamental problem. They were a nice algorithm implementation and the application was terrible. People didn't use them. *Because of the ability to put it centrally*, as opposed to having to replicate the hardware everywhere, you were afforded a cost savings per user that you couldn't even begin to approach *in a phone . . .*

Once we got into being able to *make it a piece of the switch or of the network*, you could spend lots of money on hardware because it got distributed across tens of thousands of customers. So we were able to now really develop something people could use from a cost perspective . . .

I mean, even then, phones were being given away. People wouldn't spend any more for a peripheral to a phone, even to save their lives, until afterwards. I mean, it's just like today. You go sign up for cellular service, you get a phone for free or something like that. And these things were coming in at \$300 retail. You know, that was insane. So nobody -- didn't take off. It was a failure as a commercial product. And as a speech recognizer.

¹⁴ The inventors are all Voice Control Systems' employees. The company was eventually sold to Phillips Electronics, N.V., and certain Phillips' assets were thereafter sold to ScanSoft.

¹⁵ Foster testified that the NEC phone was a mobile phone-based product sold in Europe that had the same functionality as the Uniden VoiceDial product. Foster Dep., p. 123, 58-60.

Foster Dep. at 118-120.

Foster testified that he approached McCaw Cellular through a joint venture, named MetroCell. McCaw and MetroCell were cellular network service providers. They agreed to purchase, and did purchase, “an intelligent peripheral . . . in other words a box that sat next to the switch . . . to accomplish voice dialing.” Foster Dep. pp. 76-78. Foster explained that the technology described in the '966 patent is the network-based technology that Voice Control Systems marketed to McCaw and other cellular network service providers:

Q. Do you remember what triggered it [the decision to file the application to which the '966 patent claims priority]?

A. Yeah. It was triggered by me believing we were going to get somewhere marketing to the wireless folks. In other words, that there was going to be a product out there that was generally known. And knowing where we are and what we had to do to get there, I wanted some way to protect it. . . .

* * *

Q. And when you say -- I think you said it was triggered by you were going to get somewhere in the market with the wireless folk, is that --

A. Service providers.

Q. The wireless service providers?

A. Uh-huh.

Foster Dep. p. 110. *See also* Foster Dep. at p. 80.

Voice Control Systems' handset technology had been disclosed in publications in 1986 and was sold by Uniden in the late 1980s. Foster Dep. 54,127-128; Tutorial Ex. E (1989). The application to which the '966 patent claims priority was filed in April, 1992. Foster explained that, when he approached patent counsel, the earlier handset-based voice dialer could no longer

be patented. However, the system developed for McCaw -- the "centrally" located network-based facility -- could, and, he concluded, should, be patented. Foster testified:

Q. When approximately did the three of you [the inventors Foster, Schalk and Bareis] begin this collaborative process which resulted in the invention that's described in this family of patents [the '966 patent and three other patents claiming priority to the same application]?

A. I don't really remember.

Q. Do you remember if I -- the first patent application was filed in April of 1992. Can you tell me, we started two years before that, five years before that? I assume it's a number of years to develop this technology, but I may be wrong.

A. Yeah, this invention -- let me answer it this way: We got together with patent counsel and learned about patents because none of us were experts in it. And had to go through, jump through some steps with the patent attorney to determine what inventions were not accomplished prior or, you know, too early in the process. I don't even remember how long that was. I think it was a year. I might be wrong.

But we went through things like we had to produce marketing documents to show him of all our products and -- maybe not all of them, but anything relevant, you know, so that we could sort of bound this and determine where we could start with the invention.

It was my opinion we had patentable technology of other ilks, other inventions, let's say, that we chose not to patent earlier. So there's this continuum of what we're doing.

But it centers around the inventions that we made that went into the McCaw cellular voice dialing system; that was the trigger that caused us to say, well, this stuff is now getting out. That was a commercial product, you know, something that the average person I think would think is commercial. We sold it. They put it into general use among their wireless customers in Dallas, Texas.

Q. Okay.

A. *So we, you know, we thought we'd patent it. We went back and looked at any number of inventions that we had, figured out the ones that were possible or let's say, the aspects of this that were possible to protect with patent, with a patent law and dealt with*

those in these patents. And the core of it is, you know, voice dialing in this wireless environment.

Foster Dep. pp. 24-26 (emphasis supplied), *see also* p. 109.

Foster's testimony demonstrates that the inventors' intention was to obtain patents that covered a centrally-located voice dialing product and method that allowed wireless service providers to offer a voice dialing service to wireless customers -- not the handset-based voice dialing product and method that could no longer be patented. ScanSoft's proposed claim construction is inconsistent with that objective.

Foster's testimony was augmented by the testimony of Thomas Schalk, another inventor of the '966 patent. Schalk distinguished between "the cellular switch-based recognizer" which he called an "off-board" recognizer and an "embedded" (or "on-board") voice recognizer that is "part of the phone." Schalk Dep. pp. 118-119. Schalk explained that in an "embedded" product "you have a direct connection between what you speak into and the recognizer; whereas with an off-board you are going through some wireless media to communicate" with the voice recognizer "over the air waves." Schalk Dep. pp. 118-120. Schalk was asked -- in precisely the words of Claim 1 of the '966 patent -- if the method used by the Uniden mobile-phone-based product was "a speech recognition method for a mobile telecommunications system." He responded in the negative, and explained his reasoning:

Q. . . . Is it correct that the method incorporated in the Uniden system was a speech recognition method for a mobile telecommunications system?

MR. SCHECTER: Objection.

THE DEPONENT: Not a mobile communications system. It was an interface to the phone, to the dialing process of the phone.

Therefore, ScanSoft's present attempt to extend the language of Claim 1 of the '966 patent to encompass a voice recognition functionality that is entirely internal to a single cell phone is flatly inconsistent with the meaning of the words of the claim, as understood by an inventor of the patent.

The Foster and Schalk testimony offer an unusually revealing glimpse at the real world. Their testimony demonstrates that ScanSoft's proposed claim construction is a cynical effort to contort the claims of the '966 patent into something that the inventors themselves never intended.

III. What Is A Command, A First Command Type And A Second Command Type?

| Claim Limitation | Voice Signal's Proposed Construction | ScanSoft's Proposed Construction |
|---|---|--|
| Receiving a command . . . ; ...Determining whether the command is a first or second command type | <p>The mobile telecommunications system, which includes a voice recognizer;</p> <p>(1) must receive sounds spoken by a mobile telecommunications user that it recognizes as one of a pre-established set of command words. A command word tells the voice recognizer what it should do next; and</p> <p>(2) must determine whether the command word that it has recognized indicates the user will next speak a series of numbers (digits) or a keyword. A word or phrase that is a first command type must, therefore, be different from a word or phrase that is a second command type.</p> | <p>A process performed by the speech recognition system to recognize whether the user has spoken a first or second command type so that the system then can perform the routines associated with the first or second command type.</p> <p>A "command" is a word, phrase, or numeric digit, alone or in combination with other words, phrases and/or digits, used to direct the speech recognition system to take some action.</p> <p>A "first command type" is one category of commands recognized by the speech recognition system.</p> <p>A "second command type" is another category of commands recognized by the speech recognition system.</p> |

Claim 1 of the '966 patent requires the mobile telecommunications system (1) receive a command from a user, and (2) to determine whether that command is a first command type or a second command type. If the command is a first command type, the system must collect digits representing a telephone number to be dialed. If the command is a second command type, the system must determine whether a previously stored phone number is associated with a keyword received from the user.

On its face, this appears to claim the type of ordered speech (“syntax”) that is described in the '966 patent. The user says one word (or phrase) -- a first command type -- if he intends thereafter to speak a telephone number. He says a different word (or phrase) -- a second command type -- if he intends thereafter to speak a keyword (a word that substitutes for a phone number). 6:23-36. If the user says a word/phrase that the system recognizes as a first command type, the system “listens” for and collects the digits of a phone number. 6:46-51. If the user speaks a different word/phrase that the system recognizes as a second command type, it “listens for” a keyword and determines whether a phone number has previously been associated with that keyword by the user. 7:34-37, 7:60-8:10.

This interpretation is also consistent with the words of the claim. The claim distinguishes between commands, on the one hand, and characters, on the other. (The claim requires a voice recognizer that is “capable of recognizing commands and characters.”) “Characters” are presumably what the claim also refers to as “digits” in a telephone number. Thus, in the parlance of the claim, the words “*dial*”, and later “*617-248-5207*,” are a first command type (“*dial*”), plus “digits” representing a phone number. Similarly, the words “*call*”, and later “*home*,” consist of a second command type (*call*) and a keyword (*home*). By referring to first and second type commands, the claim allows other words or phrases to substitute for *dial* and *call*, but a first

command type must be a word or phrase that is different from the word or phrase that constitutes the second command type -- otherwise there would be no reason to call one a first type, and the other a second type, command.

This interpretation is consistent with the way these terms are used in the written description. A command tells the system what to do next. Spoken telephone numbers and keywords tell the system, directly or indirectly, what telephone number to dial. In the written description, the words *Dial*, *Call* and *Directory* -- which tell the system what to do next (listen for a telephone number, listen for a keyword, or listen for a name) -- are consistently called commands.¹⁶ *E.g.*, 6:25, 27, 30, 31, 34, 35, 7:36, 8:31. By contrast, the numbers in a spoken telephone number are referred to as “digits,” not “commands.” 6:51, 7:6, 10. And, the specification calls location and relationship identifiers (*home, office, secretary, friend, information*) “keywords,” not “commands.”

Moreover, the '966 patent teaches the use of different commands -- *Dial* as compared to *Call* -- to trigger different user interface “routines.” 6:23-35, Fig. 5. The receipt and recognition of a first command type (*e.g., Dial*) triggers a routine that includes collecting digits that represent a phone number. The receipt and recognition of a different, second command type (*e.g., Call*) triggers a routine that includes the step of determining whether a phone number is associated in the system’s memory with a keyword spoken by the user. *Compare* Fig. 6 to Fig. 7.

Voice Signal’s proposed claim construction is consistent with the testimony of ScanSoft’s Rule 30(b)(6) witness on the subject of the meaning of the claims of the '966 patent. ScanSoft’s designated witness was Michael Phillips, its Chief Technology Officer, and a clear partisan. Phillips testified that first and second command types are words or phrases that proceed

¹⁶ Other words that tell the system what to do next, *e.g.*, “verify,” “store” and “send,” are also called “commands” in the written description, but those commands are not relevant to Claim 1.

the speaking of a telephone number, or a keyword, respectively. The command word/phrase is different if the speaker intends next to speak digits from the command word/phrase that is spoken if the speaker intends next to say a keyword. The use of different command words allows the system to “more carefully tailor both recognition vocabularies and even user interfaces [e.g., prompts] to the type of command the user is trying to provide.” Phillips Dep. pp. 52-57. *See also* pp. 121-122 (stating that Claim 1 of the '966 patent requires “a command to select among digit dialing and name dialing”).

ScanSoft now argues that the entire word series *call 617-248-5207* or *call home* is a command, and that the claim would encompass a command structure that did not require that different words/phrases be spoken before a telephone number or a keyword is spoken.¹⁷ ScanSoft’s argument is (a) inconsistent with the usage of the word “command” in the specification, (b) inconsistent with the conceptual teaching of the patent (that different commands precede and trigger a different series of prompts and user inputs), and (c) inconsistent with the testimony of its own Chief Technology Officer and designated Rule 30(b)(6) witness.

IV. Collecting Digits Representing A Phone Number.

| Claim Term | Voice Signal’s Proposed Construction | ScanSoft’s Proposed Construction |
|---|--|--|
| If the command is the first command type, <i>collecting digits representing a telephone number to be dialed received from the mobile telecommunications user.</i> | The mobile telecommunications system, which includes a voice recognizer, must collect a series of spoken numbers that area telephone number. | The speech recognition collects a series of numbers that the system recognizes as a complete telephone number. |

One of the steps of Claim 1 of the '966 patent requires that the mobile telecommunications system collect “digits representing a telephone number to be dialed.” This

¹⁷ ScanSoft is driven to this position because it has determined that, while early versions of the accused product require the use of different commands -- “digit dial,” before speaking a phone number, and “name dial,” before speaking a name -- later versions of the accused product do not.

phrase means what it says. The system must collect the digits that make up a telephone number - the telephone number that the user spoke and the system has thus “received from” the user. The claim does not specify how the system should collect the digits representing a phone number. It merely requires that digits be collected that are a telephone number. ScanSoft would engraft on the claim language the further limitation that the system have intelligence that would cause it to “know” that a complete telephone number has, in fact, been spoken. ScanSoft Mem. at 19-20.

The plain meaning of the words of the claim is the system must collect the group of digits that are the telephone number that the user wishes to dial. The claim limitation is met -- the system “collects the digits representing a telephone number” -- any time a user speaks a series of digits that are, in fact, a telephone number. The claim simply does not say that the system must “recognize the spoken digits as a telephone number.” If, on a particular occasion, a user made a mistake and spoke a series of digits that were not a telephone number, there would, on that particular occasion, be no infringement.

ScanSoft asks the Court to read a limitation into the claim from the specification. “A limitation that does not exist in a claim should not be read into that claim.” *Biovail Corp. v. Andix Pharmaceuticals, Inc.*, 239 F.3d 1297, 1301 (Fed. Cir. 2001). that is not apparent from the claim language, and is based upon a one-sentence disclosure (7:5-7) in the patent. Its argument must fail because the court may not read a limitation into the claim from a specification.

Moreover, the specification does not even teach the claim limitation that ScanSoft proposes. ScanSoft relies on a one-sentence statement in the '966 patent (7:5-7) that discloses only that the system stops listening for digits once it has recognized the total number of digits that it “expects” to hear (presumably because it has been programmed to stop listening after it

has recognized a defined number of digits). The system does not otherwise “recognize” the numbers as a phone number. In this respect, the '966 disclosure is like other prior art systems, each of which employed a method that caused the voice recognizer to stop listening for digits after a string of digits was spoken. These systems stopped listening either because they were programmed to recognize a word that “told” the system to stop listening for digits (*e.g.*, the prior art Voice Control Systems/Uniden product, Tutorial, Ex. E, p. 4), or because they were programmed to stop listening when they detected a pause in speech of a predefined length (*e.g.*, the prior art Texas Instruments product, Tutorial, Ex. F, p. 97) or because they were preprogrammed to listen for a specific number of digits (*e.g.*, the prior art Pirz patent, Tutorial, Ex. A, 7:5-11). The system of the '966 patent did not “recognize” that a digit string was a phone number any more than any of these prior art systems.¹⁸

In sum, the claim language is plain. The claim requires only that the digits representing a phone number be collected, one way or another. The claim language does not specify a particular method that must be used or require that other intelligence be provided. ScanSoft may not import a limitation from the specification into the claim. And, in any event, the '966 specification does not teach the limitation for which ScanSoft advocates.¹⁹

¹⁸ A system that is programmed to “listen” for a predefined number of digits does not “know” whether those digits are a telephone number. For example, the digit string 1-111-111-1111 has the appropriate number of digits, but it is not an actual telephone number.

¹⁹ The claim limitation “determining whether a previously stored phone number is associated with a keyword” is not in dispute. It means that the speech recognition system determines whether there is, in its memory, a match between the keyword spoken by the user and a telephone number previously stored in the system and previously associated with that keyword.

CONCLUSION

For the reasons stated above, the claim construction proposed by Voice Signal should be adopted.

Respectfully submitted,

VOICE SIGNAL TECHNOLOGIES, INC.

By its attorneys,

/s/ Wendy S. Plotkin

Robert S. Frank, Jr. (BBO No. 177240)
Sarah Chapin Columbia (BBO No. 550155)
Paul D. Popeo (BBO No. 567727)
Paul E. Bonanno (BBO No. 646838)
Wendy S. Plotkin (BBO No. 647716)
CHOATE, HALL & STEWART
Exchange Place
53 State Street
Boston, MA 02109
(617) 248-5000

Dated: June 3, 2005